

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2003, 01:52:07 ; Search time 3095 Seconds
(without alignments)
4137.394 Million cell updates/sec

Title: US-09-847-081B-2

Perfect score: 2270

Sequence: 1 MSMSVALLWVSTSEVSNQ.....IYAKSLVPPNRTSSPLAKT 440

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14531402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09847081/runat_01042003.120128.26739/app_query.fasta_1.583
-DB=GenEmbl -OFMT=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09847081 -CGN_1.1_3745.runat_01042003.120128.26739 -NCPU=6 -ICPU=3
-NO_XLPY -NO_MMAL -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_em.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match % | Length | DB ID | Description |
|------------|--------|---------|--------|-------------|---------------------|
| 1 | 1827 | 80.5 | 1295 | 8 CAPSY1 | X68017 C. annuum ps |
| 2 | 1751 | 77.1 | 1786 | 8 TOMCBE | M84744 Tomato phyt |
| 3 | 1750 | 77.1 | 1239 | 6 A68204 | A68204 Sequence 2 |
| 4 | 1750 | 77.1 | 1239 | 6 AR156064 | AR156064 Sequence |
| 5 | 1750 | 77.1 | 1239 | 6 BD005486 | BD005486 Enhanceme |
| 6 | 1744 | 76.8 | 1239 | 6 A68203 | A68203 Sequence 1 |
| 7 | 1744 | 76.8 | 1239 | 6 AR156063 | AR156063 Sequence |
| 8 | 1744 | 76.8 | 1239 | 6 BD005485 | BD005485 Enhanceme |
| 9 | 1733 | 76.3 | 1614 | 8 LER1PE | Y00521 Tomato frui |
| 10 | 1732 | 76.3 | 1646 | 6 A21360 | A21360 L. esculentu |
| 11 | 1732 | 76.3 | 1646 | 6 AR007503 | AR007503 Sequence |
| 12 | 1668 | 73.5 | 1355 | 8 LERYCTOM5 | X67144 L. esculentu |
| 13 | 1626 | 71.6 | 1506 | 8 AF220218 | AF220218 Citrus un |
| 14 | 1626 | 71.6 | 1689 | 8 AB037975 | AB037975 Citrus un |
| 15 | 1606 | 70.7 | 1773 | 8 AF152892 | AF152892 Citrus x |
| 16 | 1598 | 70.4 | 1637 | 8 HAN308385 | AJ308385 Helianthu |
| 17 | 1595 | 70.3 | 2868 | 6 E15683 | E15683 Gentiana lu |
| 18 | 1590 | 70.0 | 1598 | 8 HAN304825 | AJ304825 Helianthu |
| 19 | 1580.5 | 69.6 | 1376 | 8 AY099482 | AY099482 Tagetes e |
| 20 | 1568.5 | 69.1 | 1576 | 8 CMPSYIMR | Z37543 C. melo PSY1 |
| 21 | 1568 | 69.1 | 1371 | 8 AF251015 | AF251015 Tagetes e |
| 22 | 1559.5 | 68.7 | 1372 | 6 A48337 | A48337 Sequence 1 |
| 23 | 1557 | 68.6 | 1921 | 6 E15681 | E15681 Gentiana lu |
| 24 | 1544.5 | 68.0 | 1853 | 8 AY058287 | AY058287 Arabidops |
| 25 | 1542 | 67.9 | 1703 | 8 AY085565 | AY085565 Arabidops |
| 26 | 1531 | 67.4 | 2085 | 6 E15680 | E15680 Gentiana lu |
| 27 | 1530 | 67.4 | 1119 | 8 TOMPSY2A | L23424 Lycopersico |
| 28 | 1526.5 | 67.2 | 1769 | 8 ATHPHYSYN | L25812 Arabidops |
| 29 | 1523 | 67.1 | 1932 | 6 E15682 | E15682 Gentiana lu |
| 30 | 1519 | 66.9 | 1548 | 8 NPPSY | X78814 N. pseudonar |
| 31 | 1495.5 | 65.9 | 4073 | 8 LEPSPS | X60440 L. esculentu |
| 32 | 1431.5 | 63.1 | 1304 | 6 AX350858 | AX350858 Sequence |
| 33 | 1334.5 | 58.8 | 1538 | 8 AB032797 | AB032797 Daucus ca |
| 34 | 1282.5 | 56.5 | 3707 | 8 LEGTOM5 | X60441 L. esculentu |
| 35 | 1268.5 | 55.9 | 3200 | 8 AF009954 | AF009954 Arabidops |
| 36 | 1268.5 | 55.9 | 75188 | 8 AB005238 | AB005238 Arabidops |
| 37 | 1212 | 53.4 | 1060 | 8 AY024350 | AY024350 Oryza sat |
| 38 | 1064 | 46.9 | 681 | 8 AF196963 | AF196963 Bixa orel |
| 39 | 1059 | 46.7 | 686 | 8 AF158024 | AF158024 Tagetes e |
| 40 | 1054.5 | 46.5 | 1278 | 8 DBU91900 | U91900 Dunaliella |
| 41 | 1052 | 46.3 | 1859 | 8 AF305430 | AF305430 Haematoco |
| 42 | 1026.5 | 45.2 | 139380 | 2 CNS08CAG | AL831803 Oryza sat |
| 43 | 1015 | 44.7 | 5995 | 8 ZM032636 | U32636 Zea mays ph |
| 44 | 1010 | 44.5 | 3222 | 8 AY024351 | AY024351 Oryza sat |
| 45 | 901 | 39.7 | 1465 | 1 SSPHYSYN | X69172 Synechocyst |

ALIGNMENTS

RESULT 1

A68204
LOCUS A68204 1239 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 2 from Patent WO9746690.
ACCESSION A68204
VERSION A68204.1 GI:4759372
KEYWORDS
SOURCE tomato.
ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
REFERENCE 1 (bases 1 to 1239)
AUTHORS Drake, C.R., Bird, C.R. and Schuch, W.W.
TITLE ENHANCEMENT OF GENE EXPRESSION
JOURNAL Patent: WO 9746690-A 2 11-DEC-1997;
 ZENECA LTD (GB)
FEATURES Location/Qualifiers
 source 1..1239
 /organism="Lycopersicon esculentum"
 /db_xref="taxon:4081"
 /clone="GTON5" PHYTOENE SYNTHASE GENE"
BASE COUNT 377 a 207 c 317 g 338 t
ORIGIN
 Alignment Scores:
 Pred. No.: 3,42e-142 Length: 1239
 Score: 1750.00 Matches: 344
 Percent Similarity: 86.84% Conservative: 32
 Best Local Similarity: 79.45% Mismatches: 33
 Query Match: 77.09% Indels: 24
 DB: 6 Gaps: 3
US-09-847-081b-2 (1-440) x A68204 (1-1239)
 QY 3 MetSerValAlaLeuLeuTrpValSerProThrSerGluValSerAsnGlyThrGly 22
 |||||
 Db 1 ATGTCTGTCCTTGTATGGTGTGTTCTCTCT---TGTGACGCTCAATGGGACAAGT 57
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
 |||||
 Db 58 TTATGGAATCATGTCGGGAGGAAACCGTTTTTTTGTATTCATCG-----AGG 105
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgGlyLeuLysGlyGlyArgGlnArgTrpAsnPhe 62
 |||||
 Db 106 CATAGGAATTTGTGTCATGAGAGATCAATAGAGT-----144
 QY 63 GlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
 |||||
 Db 145 -----GGTGAAGCAAACTAATAATGA 168
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
 |||||
 Db 169 CGGAATTTCTGTACGGTCTGTATTTGGCTACTCTCGAGAACGCGATGACA 228
 QY 103 SerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuValLysArgGlnLeu 122
 |||||
 Db 229 TCGGAACAGATGCTCATGATGTGGTTTGAGCAGCAGCCCTTGTGAGAGCGCACTG 288
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142
 |||||
 Db 289 AGATCTACCAATGAGTAGAAGTGAAGCGGATATACCTATTCCGGGGAATTTGGCGTTG 348
 QY 143 LeuSerGluAlaLysArgPheCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
 |||||
 Db 349 TTGAGTGAACCATATGATAGGTGTGGTGAAGTATGTGCAGAGTATGCAAGACGTTTAAAC 408
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrValTrp 182
 |||||
 Db 409 TTAGGAACATGCTATGACTCCCGAGAGAAAGGCGCTATCTGGGCAATATATGATGG 468
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
 |||||

Db 469 TGCAGAGAACACAGATGAACCTTGTGTGATGGCCCAACCCATCATATATATACCCCGCAGCC 528
 QY 203 LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222
 |||||
 Db 529 TTAGATAGGTGGGAAAATAGCTAGAGATGTTTTCATGGCGGCCCATTTGACATGCTC 588
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242
 |||||
 Db 589 GATGGTGGCTTTGTCGATACAGATTTCTTAACCTTCCAGTGTATATTCAGCCATTGACAGAT 648
 QY 243 MetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeu 262
 |||||
 Db 649 ATGATTGAGGAATGCGTATGGACTTGAGAAAATCGAGATACAAAACCTTCGAGCAACTA 708
 QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
 |||||
 Db 709 TACCTTTATTTATTATGTTGCTGTGACGTGGTGGTGTGATGAGTGTTCCTCAATATGGGT 768
 QY 283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302
 |||||
 Db 769 ATCGCCCTGAATCAAGGCAACACAGAGCGGTATATATGCTGCTTTGGCTCTGGGG 828
 QY 303 LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgGlyArg 322
 |||||
 Db 829 ATCGCAATCAATTAACATACATCTCAGAGATGTTGGAGAAGATGCCAGAGAGGAGA 888
 QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342
 |||||
 Db 889 GTCTACTTGGCTCAAGATGAATAGCACAGCGAGGTCTATCCGATGAAGATATATTTGCT 948
 QY 343 GlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys 362
 |||||
 Db 949 GGAAGGGTGACCCATAAATGGGAATCTTTATCAAGAAACAAATACATACATGGCGCAAGAAAG 1008
 QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382
 |||||
 Db 1009 TTCTTTGATGAGCAGAGAAAGCGGTGACAGAAATGAGCTCAGCTAGTAGATCCCTGTA 1068
 QY 383 LeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402
 |||||
 Db 1069 TGGGCATCTTGTGCTGTACCGCAAAATACTAGATGAGATTGAAGCCAATGACTACAAC 1128
 QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422
 |||||
 Db 1129 AACTTCACAAAGAGAGCATATGTGAGCAAAATCAAGAAAGTTGATTGCTACCTATTGCA 1188
 QY 423 TyrAlaLysSerLeuValProProAsnArgThrSerSer 435
 |||||
 Db 1189 TATGCAAAATCTCTTGTGCTCTCTACAAAACCTGCCTCT 1227
 RESULT 4
 ARI56064
 LOCUS ARI56064 1239 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 2 from patent US 6239331.
 ACCESSION ARI56064
 VERSION ARI56064.1 GI:15124117
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1239)
 AUTHORS Drake, C.Rachel., Bird, C.Roger, and Schuch, W.Walter.
 TITLE Enhancement of tomato phytoene synthase gene expression with a modified DNA
 JOURNAL Patent: US 6239331-A 2 29-MAY-2001;
 FEATURES Location/Qualifiers
 source 1..1239
 /organism="unknown"
 BASE COUNT 377 a 207 c 317 g 338 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,42e-142 Length: 1239
 Scores: 377 a 207 c 317 g 338 t

[illegible]

```
QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
Db 58 TTCTGGAATACATCCGGGAGGAAACCGCTTTTGGTTCATCG-----AGG 105
QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysGlyGlyArgGlnArgTrpAsnPhe 62
Db 106 CATAGGANTTGGTCCCATGAGAGATCAATAGAGT-----144
QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
Db 145 -----GGTGGAAAGCAACAAATAATATGGA 168
QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
Db 169 CGGAATATTTCTAGCGGTGCTATATTTGGCTACTCATCTGGAGAACGGAGATGACA 228
QY 103 SerGluLysLysValTrpAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
Db 229 TCGGACAGATGCTATGATGTTGGTTTGAGCGCAGCGCTTGTGTGAGAGGCAACGTG 288
QY 123 ArgSerThrAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142
Db 289 AGATCTACCAATCAGTTAGAGTGAAGCGGATATACCTATTCCGGGGAATTTGGCGTTG 348
QY 143 LeuSerGluAlaLysArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
Db 349 TTGAGTGAAGCATATGATAGTGTGGTGAAGTATGTGCAGAGTATGCAAGAGCGTTTAA 408
QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrValTrp 182
Db 409 TTAGGAACATATGATATGATCTCCGAGAGAAAGGCGCATCTCGGCAATATATGATGG 468
QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
Db 469 TGCAGAAAGACATCACTGTTGATGGCCCAACGCATCATATATACCCCGGCAGCC 528
QY 203 LeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222
Db 529 TTAGATAGGTGGAAATAGGCTAGAGATGTTTTCATGGGGCGGCAATTTGACATGCTC 588
QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242
Db 589 GATGGTGTCTTCCCATACAGTTTCTACTTCCAGTTGATATTCAGCCATTCAGAGAT 648
QY 243 MetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGluLeu 262
Db 649 ATGATTGAAGAAATGCGTATGGACTTGAGAAATCGAGATACAAAACTCGACGAACTA 708
QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
Db 709 TACCTTTATTTATTTATTTGCTGTGTGCTAGCGTTGGTTGGTTGATGAGTGTTCCTAATTTATGGGT 768
QY 283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302
Db 769 ATGCCCTTGAAATCAAGGCAACACAGAGCGGTATATATCTGCTTTGGCTCTGGGG 828
QY 303 LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArg 322
Db 829 ATCGCAAAATCAATTAATACTACTACAGATGTTGGAGAGATGCCAGAGAGAGAAGA 888
QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342
Db 889 GTCTACTTGCCTCAAGATGAATAGACAGCGAGGTCTATCCGATGAAGATATATTTGCT 948
QY 343 GlyArgValThrAspLysTrpArgAsnPheMetLysGlyGlnIleGlnArgAlaArgLys 362
Db 949 GGAAGGTGACCCATAAATGGAGAAATCTTTATGAAGAAACAAATACATAGGGCAAGAAAG 1008
QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal 382
Db 1009 TTCTTTGATGAGCGACAGAAAGCGGTGACAGAAATTCAGCTCAGCTAGTAGATTCCTCTGTA 1068
```

```
QY 383 LeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402
Db 1069 TGGGCATCTTTGGTCTGTACCCCAAAATACTAGATGAGATTGAAGCAATGACTACAAC 1128
QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422
Db 1129 AACTTCACAAAGAGACATATGTGACAAATCAAGAAGTTGATTGCATTACCTATTGCA 1188
QY 423 TyrAlaLysSerLeuValProProAsnArgThrSerSer 435
Db 1189 TATGCAAAATCTCTTGTGCTCTCTCAAAAACACTGCCTCT 1227
RESULT 6
A68203
LOCUS A68203 1239 bp DNA linear PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9746690.
ACCESSION A68203
VERSION A68203.1 GI:4759371
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE 1 (bases 1 to 1239)
AUTHORS Drake,C.R., Bird,C.R. and Schuch,W.W.
TITLE ENHANCEMENT OF GENE EXPRESSION
JOURNAL Patent: WO 9746690-A 1 11-DEC-1997;
ZENECA LTD (GB)
FEATURES
Location/Qualifiers
Source
1..1239
/organism="unidentified"
/isolate="SYNTHETIC DNA"
/db_xref="taxon:32644"
BASE COUNT 343 a 247 c 320 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 1,13e-141 Length: 1239
Score: 1744.00 Matches: 343
Percent Similarity: 86.61% Conservative: 32
Best Local Similarity: 79.21% Mismatches: 34
Query Match: 76.83% Indels: 24
DB: 6 Gaps: 3
US-09-847-081B-2 (1-440) x A68203 (1-1239)
QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
Db 1 ATGAGCGTGGCACTCTTTGGGTGGTGAGCCCA---TGCATGTGAGTAACGGCAGCTCA 57
QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
Db 58 TTTATGGAGAGTGTGAGAGAGGTAATAGATTCTTCGACAGTTCT-----CGT 105
QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysGlyGlyArgGlnArgTrpAsnPhe 62
Db 106 CACCGTAACTTGTGTAGTAAACAGCAACGTATATAACAGGGGA-----144
QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
Db 145 -----GGAGTTAAACAGCAACACCGT 168
QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
Db 169 AGAAAGTTCTCAGTTAGATCAGCAATCTTGCACACCTTAGCGGTGAGAGCACTATGACT 228
QY 103 SerGluLysLysValTrpAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
Db 229 AGCGAGCAATATGGTGTACGACGCTGCTACTTCTGTCAGAGTGCAGTAGTTAAACGTCAGTTA 288
QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142
Db 289 CGTAGTACTAAACGAACTTGAGGTAAACCTGACATTCCCAATACCTTGGACCTT 348
```

QY 143 LeuSerGluAlaThrAspArgCysGlyGluValCysAlaGluThrAlaLysThrPheTyr 162
 DB 349 CTTCTGAGCGTACACAGATCGGAGAGGTTGCGAGAAATACGCTAAACCTTCAT 408
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrAlaIleTyrValTyr 182
 DB 409 TTGGGTACCATGTTGATGACACCAAGGCGTCTGTCATATGGCTATTTACGTTGG 468
 QY 183 CysArgArgThrAspGluValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
 DB 469 TGTAGCGTACTGACGAGTTAGTGGAGCGACCTAATGCTAGTATACATACACCGCTGT 528
 QY 203 LeuAspArgThrGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222
 DB 529 CTTGACAGATGGGAGAACCGTTGGAGGACGTTTAAACGCGACCTTTTCATATGTTG 588
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgasp 242
 DB 589 GACGGAGCACTTAGTGACACTGTAGCAATTTCCCTGTGGACATCCACCTTTTCGGGAC 648
 QY 243 MetIleGluGlyMetArgMetAspLeuThrLysSerArgTyrLysThrPheAspGluLeu 262
 DB 649 ATGATCGAGGCGATGAGATGATCTCGTAAGTCTCGTTATAAGAAATTTTCATGATG 708
 QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
 DB 709 TATTTGTACTGTACTACTGTGCGAGAACCGTGGGCGCTTATGCTAGTGCCTATCATGGA 768
 QY 283 IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly 302
 DB 769 ATGCACAGAGATTAAGCTACTACTGATCTGTTACCGCGACACTACATAGGT 828
 QY 303 LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgGlyArg 322
 DB 829 ATAGCTAACAGCTTACAAATATCTTGGAGGAGTGGGTGAGGACGACAGTGGGTCT 888
 QY 323 ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla 342
 DB 889 GTGATCTCCACAGAGAGCTGCTCAAGCTGGATTGAGTGACGAGGACATTTTCGCA 948
 QY 343 GlyArgValThrAspLysThrArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys 362
 DB 949 GTCTGTGTACAGACAGTGGAGGATTTTCATGAAACACCATTCACCGTCTCGTAAA 1008
 QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTyrProVal 382
 DB 1009 TTTTTCGAGAACTGAAAGGAGTACTGAGCTTCTAGTGCATCAAGTTTCCAGTT 1068
 QY 383 LeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402
 DB 1069 TGGGCGAGCTTGTGCTCTATAGAAGATTTTGGACGAATTCGAGGCTAACGATTATAT 1128
 QY 403 AsnPheThrArgArgAlaTyrValSerLysProLysLysLeuThrLeuThrProIleAla 422
 DB 1129 AATTITACTAAGCTGCTTACTTCTTAAGACAAACAACTTATCGCTTCCATCGCT 1188
 QY 423 TyrAlaLysSerLeuValProProAsnArgThrSerSer 435
 DB 1189 TAGCTAAGAGCTTGGTTCACCACTAAGACAGCTAGC 1227
 RESULT 7
 ARL156063
 LOCUS ARL156063 1239 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 1 from patent US 6239331.
 ACCESSION ARL156063
 VERSION ARL156063.1 GI:15124116
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1239)
 AUTHORS Drake, C. Rachel., Bird, C. Roger., and Schuch, W. Walter.
 TITLE Enhancement of tomato phytoene synthase gene expression with a

modified DNA
 JOURNAL Patent: US 6239331-A 1 29-MAY-2001:
 FEATURES Location/Qualifiers

source 1..1239
 /organism="unknown"
 BASE COUNT 343 a 247 c 320 g 329 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1..13e-141 Length: 1239
 Score: 1744.00 Matches: 343
 Percent Similarity: 86.61% Conservative: 32
 Best Local Similarity: 79.21% Mismatches: 34
 Query Match: 76.83% Indels: 24
 DB: 6 Gaps: 3

US-09-847-081B-2 (1-440) x ARL156063 (1-1239)

QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
 DB 1 ATGAGCGTGGCAGCTTCTTTGGGTGGTGAGCCCA---TGGCATGTGAGTAACGCACTTCA 57
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
 DB 58 TTTATGGAGAGTGTGAGAGAAGGTAATAGATTCTTCGACAGTTCT-----CGT 105
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTTrpAsnPhe 62
 DB 106 CACCGTAACCTTGTAGTAACGACGTATAACAGGGGA----- 144
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
 DB 145 -----GGAGGTAACAGCAACAACGCT 168
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
 DB 169 AGAAGTTCTCAGTTAGATCAGCAATCCTGTCAACACCTAGCGGTGAGAACTATGACT 228
 QY 103 SerGluValLysLysValTyrAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
 DB 229 ACGGACCAATGGTGTACGACGCTGCTACTTCGTCAAGTGCACCTAGTTAAACGTCAGTTA 288
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142
 DB 289 CGTAGTACTAACGAACCTTCAGGTTAAACCTGACATCCAACTACCTGGAACCTTGGACTT 348
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
 DB 349 CTTTCTGAGGCTTACGACAGATCGGAGAGGTTTGGCAGAAATACGCTAAACCTTCAAT 408
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrAlaIleTyrValTyr 182
 DB 409 TTGGGTACCATGTTGATGACACCAAGAGCGCTGTCGCAATATGGCTATTTACGTTGG 468
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
 DB 469 TGTAGCGTACTGACGAGTTAGTGGAGGACCTAATGCTAGTTACATAACACCGCTGCT 528
 QY 203 LeuAspArgThrGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222
 DB 529 CTTGACAGATGGGAGAACCGTTTGGAGGACGTTTAAACGCGACCTTTTCGATATGTTG 588
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgasp 242
 DB 589 GACGGAGCACTTAGTGACACTGTAGCAATTTCCCTGTGGACATCCACCTTTTCGGGAC 648
 QY 243 MetIleGluGlyMetArgMetAspLeuThrLysSerArgTyrLysThrPheAspGluLeu 262
 DB 649 ATGATCGAGGCGATGAGATGATCTCGTAAGTCTCGTTATAAGAAATTTTCATGATG 708
 QY 263 TyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly 282
 DB 709 TATTTGTACTGTACTACTGTGCGAGAACCGTGGGCGCTTATGCTAGTGCCTATCATGGA 768

| QY | 283 | IleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly | 302 |
|------------------------|--|--|----------------------------|
| DB: | US-09-847-081B-2 (1-440) x BD005485 (1-1239) | | |
| DB | 769 | ATTGCACGAGAGTAAGACTACTACTGTAATCTGTTTACACCGACGACTAGCATTAGGT | 828 |
| QY | 303 | LeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgGlyArg | 322 |
| DB | 829 | ATAGCTAACCCAGCTTACAAATATCTTGAGGAGCTGGGTGAGGACGACGCTAGGGTCTGT | 888 |
| QY | 323 | ValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla | 342 |
| DB | 889 | GTGTATCTCCACAGGACGAGCTCGCTCAAGCTGGATTGAGTCACGAGGACATTTTCGCA | 948 |
| QY | 343 | GlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys | 362 |
| DB | 949 | GGTCGTGTACAGACAGTGGAGGATTTTCATGAAAAGCAGATCACCGTGTCTGTAA | 1008 |
| QY | 363 | PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpProVal | 382 |
| DB | 1009 | TTTTTCGACCAAGCTGAAAGGAGTACTGAGCTTCTAGTCATCAAGTTTCCAGTT | 1068 |
| QY | 383 | LeuThrAlaLeuLeuTyrArgLysIleLeuAspGluLeuAlaAsnAspTyrAsn | 402 |
| DB | 1069 | TGGCCAGCCTTGTCTCTATAGAGATTTTGACGAAATCGAGCTACGATTATAT | 1128 |
| QY | 403 | AsnPheThrArgAlaTyrValSerLysProLysLeuLeuThrLeuProIleAla | 422 |
| DB | 1129 | AAATTTACTAAGCTGCTTACGTTTCTAAGAGCAAAATCTTATCGTCTCTCCAAATCGCT | 1188 |
| QY | 423 | TyrAlaLysSerLeuValProProAsnArgThrSerSer | 435 |
| DB | 1189 | TACGCTAAGAGCTGTGTTCCACCACTAAGACAGCTAGC | 1227 |
| RESULT 8 | | | |
| LOCUS | BD005485 | 1239 bp | DNA linear PAT 31-JAN-2002 |
| DEFINITION | Enhancement of gene expression. | | |
| ACCESSION | BD005485 | | |
| VERSION | BD005485.1 | GI:18633856 | |
| KEYWORDS | JP 2001501810-A/1. | | |
| SOURCE | unidentified. | | |
| ORGANISM | unclassified. | | |
| REFERENCE | 1 (bases 1 to 1239) | | |
| AUTHORS | Drake, C.R., Bird, C.R. and Schuch, W.W. | | |
| TITLE | Enhancement of gene expression | | |
| JOURNAL | Patent: JP 2001501810-A 1 13-FEB-2001; | | |
| COMMENT | ZENECA LTD | | |
| | OS Unidentified | | |
| | PN JP 2001501810-A/1 | | |
| | PD 13-FEB-2001 | | |
| | PF 23-MAY-1997 JP 1998500302 | | |
| | PR 07-JUN-1996 GB 9611981.3 | | |
| | PI CAROLINE RACHEL DRAKE, COLIN ROGER BIRD, WOLFGANG WALTER SCHUCH | | |
| | PC C12N15/67, C12N15/82, C12N15/29, C07K14/415 | | |
| | CC Strandedness: Double; | | |
| | CC Topology: Linear; | | |
| | FE Key | | |
| | FT Location/Qualifiers | | |
| | FT source 1. .1239 | | |
| FEATURES | Location/Qualifiers | | |
| source | 1. .1239 | | |
| | /organism="Unidentified". | | |
| | /organism="unidentified" | | |
| | /db_xref="taxon:32644" | | |
| BASE COUNT | 343 a 247 c 320 g 329 t | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 1,13e-141 | Length: | 1239 |
| Score: | 1744.00 | Matches: | 343 |
| Percent Similarity: | 86.61% | Conservative: | 32 |
| Best Local Similarity: | 79.21% | Mismatches: | 34 |
| Query Match: | 76.83% | Indels: | 24 |

6

Gaps:

3


```

Db 949 GGTCTGTACAGCAAGTGGAGGATTTTCATGAAAGAGCAGATTCACCGTCTCGTAAA 1008
QY 363 PhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTyrProVal 382
Db 1009 TTTTTCGAGAGCTGAAAGGAGTACTAGCTTTCTAGTCATCAAGTTTCCAGTT 1068
QY 383 LeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn 402
Db 1069 TGGGCCAGCCTGTGCTCTATAGAAGATTTTGGACGAATCGAGGCTAAGGATTATAAT 1128
QY 403 AsnPheThrArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla 422
Db 1129 AATTTTACTAAGCGTCTAGCTTTCTAAGAGCAAAACATATGCTTTCCATCGCT 1188
QY 423 TyrAlaLysSerLeuValProAsnArgThrSerSer 435
Db 1189 TACGCTAAGAGCTTGTTCACCACTAAGACAGCTAGC 1227

RESULT 9
LERIPE
LOCUS LERIPE 1614 bp mRNA linear PLN 11-MAY-1995
DEFINITION Tomato fruit ripening specific mRNA.
ACCESSION Y00521
VERSION Y00521.1 GI:19340
KEYWORDS unidentified reading frame.
SOURCE Lycopersicon esculentum.
ORGANISM Lycopersicon esculentum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 1614)
REFERENCE
AUTHORS Bird,C.R.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1987) Bird C.R., Imperial chemical industries
PLC, plant biotechnology group, Po Box 11, The Health, Runcorn,
Cheshire, WA7 4QE, U.K.
2 (bases 1 to 1614)
REFERENCE
AUTHORS Ray,J., Bird,C., Maunders,M., Grierson,D. and Schuch,W.
TITLE Sequence of pTOM5, a ripening related cDNA from tomato
JOURNAL Nucleic Acids Res. 15 (24), 10587 (1987)
MEDLINE 88096591
PUBMED 3637037
FEATURES
source
1. .1614
/organism="Lycopersicon esculentum"
/strain="Mill."
/cultivar="Alisa Craig"
/db_xref="taxon:4081"
/clone="pTOM5"
201. .1439
/note="tomato fruit ripening specific URf"
/codon_start=1
/protein_id="CAA68575.1"
/db_xref="GI:19341"
/db_xref="SWISS-PROT:P08196"
INRGKGTNNRGKFSVRSAILATPSCGRTMTSEQMYDVVLRQALVKKQLRSTNEL
EVKPDIPGNLIGLLSEAYDRCEVCAEYAKTNLGNLMTPERRALWVWCRRT
DELVDGDNKRSYITPAALDRWENRLEDFNGFPMDLGLASDVTVSFPVDIQPFDM
EGMDRLKSYRNFDLYLYCYVAVGTGLMSYPINGIAPESKATTSYFNNAALIG
TANGLILRDVGDDARRGVYLPQDELQAGLSDEDFAGRVTDKVRIFMKKQIHR
RKFFDEAKGVTELSASRFVPSWASLVLYRKILDEIANDYNNFTKRAYVSKSQVDC
ITYCICIKASCASYKNASIQR"
246. .254
/misc_feature
/note="putative glycosylation site"
702. .710
/misc_feature
/note="putative glycosylation site"
1329. .1337
/misc_feature
/note="putative glycosylation site"
1419. .1427
/misc_feature
/note="putative glycosylation site"

```

```

BASE COUNT 515 a 241 c 388 g 470 t
ORIGIN
Alignment Scores:
Pred. No.: 1.43e-140 Length: 1614
Score: 1733.00 Matches: 346
Percent Similarity: 85.68% Conservative: 31
Best Local Similarity: 78.64% Mismatches: 36
Query Match: 76.34% Indels: 27
DB: 8 Gaps: 4

US-09-847-081B-2 (1-440) x LERIPE (1-1614)
QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
Db 195 CTCAGAATGCTGTTGCCTTGTATGGGTGTTTCTCCT---TGTGACGCTCAAAATGG 251
QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 252 ACAAGTTTCATGGAAATCAGTCGGGAGGGAACCGTTTTTTTGTGATTCATCG----- 302
QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTyr 60
Db 303 ---AGGCATAGCAATTTGGTCTCAATGAGAGAACTAATAGAGT----- 344
QY 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
Db 345 -----GGTGAAGCAAACTAAT 362
QY 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
Db 363 AATGGACGGAATTTTCTGTAGCGTCTGCTATTTTGGTACTCCATCTCGAGAACGGACG 422
QY 101 ValSerGluLysLysValTyrAspValValLeuLysGlnAlaLeuValLysArg 120
Db 423 ATGACATCGGAACAGATGCTATATGTGTTTGGAGCAGCAGCCCTTGTGAAGAGG 482
QY 121 GlnLeuArgSerThrAspLeuValLysProAspIleValValProGlyAsnLeu 140
Db 483 CAACTGAGATCTACCAATGAGTTAGAAGTGAAGCCGGATATACCTATTCCGGGAATTTG 542
QY 141 GlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThr 160
Db 543 GGCTTTGTCAGTGAAGCATATGATAGTGTGTGTAAGTATGTCAGAGTATGCAAAAGACG 602
QY 161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrPheAlaIleTyr 180
Db 603 TTTAACTTAGGAACCTATGCTAATGACTCCCGAGAGAAGAGGGCTATCTGGGCAATATAT 662
QY 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
Db 663 GTATGGTCAGAGAAGACAGATGAACCTGTTGATGGCCCAACGCATCATATATACCCCG 722
QY 201 GlnAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyArgPropheAsp 220
Db 723 GCAGCCTTAGATAGTGGGAAATAAGGCTAGAGAAGATGTTTCAATGGCGGCATTTGAC 782
QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProphe 240
Db 783 ATGCTCGATGGTCTTGTCCGATACAGTCTTCTAATCTCCAGTTGATATTCAGCCATTC 842
QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAsp 260
Db 843 AGAGATATGATTGAAGGAATGCTATGGACTTGGAGAAATCGAGATAACAAAACACTTCGAC 902
QY 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 903 GAACATATACCTTTATTGTTATTATGTTGCTGCTACGGTTGGTGTGATGAGTGTCCAAAT 962
QY 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaLeuAla 300
Db 963 ATGGGTATCGCCCTGAATCAAAAGGCAACACAGAGAGCGGTATATAATGCTGCTTGGCT 1022

```



```

Db 1143 TTTCGGAAGGTTGACCATAAATGGAGAATCTTTATGAAGAAACAAATACATAGGGCA 1202
QY 361 ArgLysPheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
Db 1203 AGAAAGTCTTTGATGAGCAGAGAAAGCGTGACAGATTGAGCTCAGCTAGTAGATTC 1262
QY 381 ProValLeuThrAlaLeuLeuLeuTyArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db 1263 CCTGTATGGGCATCTTTGCTCTTGACCGCAAAATACTAGATGAGATTGAAGCCAATGAC 1322
QY 401 TyrAsnAspPheThrArgArgAlaTyValSerLysProLys-LysLeuLeuThrLeuPr 420
Db 1323 TACAACACTTCACAAGAGACATATGTGACAAATCAAGCAAGTGAATGCAATACC 1382
QY 420 oilealatyAlaLysSerLeuValProProAsnArgThrSerProLeuAlaLys 439
Db 1383 TATTGCATATGCAAAATCTTTGCTCCTCT-----ACAAACTGCCCTCTCTTCAA 1434

RESULT 11
AR007503
LOCUS AR007503 1646 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5750865.
ACCESSION AR007503
VERSION AR007503.1 GI:3966987
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1646)
Unclassified.
AUTHORS Bird,C.Roger., Grierson,D. and Schuch,W.Walter.
TITLE Process for modifying the production of carotenoids in plants, and
DNA, constructs and cells therefor
JOURNAL Patent: US 5750865-A 2 12-MAY-1998;
FEATURES
Location/Qualifiers
source 1..1646
/organism="unknown"
BASE COUNT 529 a 249 c 388 g 480 t
ORIGIN

Alignment Scores:
Pred. No.: 1,78e-140 Length: 1646
Score: 1732.00 Matches: 346
Percent Similarity: 85.68% Conservative: 31
Best Local Similarity: 78.64% Mismatches: 36
Query Match: 76.30% Indels: 27
DB: 6 Gaps: 4

US-09-847-081B-2 (1-440) x AR007503 (1-1646)
QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
Db 195 CTCAGAAATCTGTTGGCTTGTATGGGTGTTCTCCT---TGTGACGCTCAAAATGGG 251
QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerArgPheLeu 40
Db 252 ACAAGTTTCATGAATCAGTCCGGGAGGAAACCGTTTTTTTGATTATCG----- 302
QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyValArgGlnArgTrp 60
Db 303 ---AGCATAGGAATTTGGTGTCCCAATGAGAAATCAATAGAGGT----- 344
QY 61 AsnPheGlySerLeuIleAlaAspProArgTySerSerLysGlyGlySerArgThrGlu 80
Db 345 -----GGTGGAAAGCAAACTAAT 362
QY 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
Db 363 AATGGACGGAATTTCTCTACGGTCTGCTATTGGTCTACTCCATCTGGAGACGGAGC 422
QY 101 ValSerSerGluLysValTyArgValValLeuLysGlnAlaLeuValLysArg 120
Db 423 ATGACATCGGAACAGATGCTCTATGTGTGGTTTGGAGCGAGCGCTTGTGTGAAGAG 482

```

```

QY 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeu 140
Db 483 CAACGTAGATCTACCAATAGTATAGAGTGAAGCCGGATATACCTATTCCGGGAATTG 542
QY 141 GlyLeuLeuSerGluAlaTyArgAspArgCysGlyGluValCysAlaGluTyAlaLysThr 160
Db 543 GCGTTGTTGAGTCAAGCATATATAGGTGCTGGTGAAGATATGTCAGAGTAGTCAAGACG 602
QY 161 PheTyLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTy 180
Db 603 TTTAACCTTAGAACATATGCTAATGCTCCCGAGAGAGAGGCTATCTGGGCAATATAT 662
QY 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
Db 663 GTATGGTGCAGAAGAACACATGAACCTGTTGATGGCCCAACCATCATATATACCCCG 722
QY 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220
Db 723 GCAGCCTTAGATAGGTGGGAAATAGGCTAGAGATGTTTCAATGGCGGCCCATTTGAC 782
QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
Db 783 ATGCTCGATGGTGGCTTGTCCGATACAGTTCTTAACCTTCCAGTTGATATTCAGCCATC 842
QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyLysThrPheAsp 260
Db 843 AGAGATATGATTGAAGGAATGGTATGGACTTGAGAAATCGACATACAAAACTTCGAC 902
QY 261 GluLeuTyLeuTyrcystTyTyValAlaGlyThrValGlyLeuMetSerValProVal 280
Db 903 GAACATATACCTTATTGTTATTATGTTGCTGGTACGTTGGTGGTGGTGGTGGTGGT 962
QY 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyTrpAsnAlaAlaLeuAla 300
Db 963 ATGGGTATGCGCCCTGAATCAAGGCAACACAGAGAGCGGTATATATGCTGCTGGCT 1022
QY 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320
Db 1023 CTGGGGATCGCAANTCAATTAACATACATCTACAGAGATGTTGGAGAGATGCCAAGA 1082
QY 321 GlyArgValTyLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
Db 1083 GCAAGAGTCTACTTGCCTCAAGATGAATTAGCACAGCAGGCTATCCGATGAAGATATA 1142
QY 341 PheAlaGlyArgValThrAspLysTyArgAsnPheMetLysLysGlnIleGlnArgAla 360
Db 1143 TTTGCTGGAAGGTTGACCGATAAATGGAGATCTTTATGAAGAAACAAATACATAGGCA 1202
QY 361 ArgLysPheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
Db 1203 AGAAAGTCTTTGATGAGCAGAGAAAGCGTGACAGAAATTTGAGCTCAGCTAGTAGATTC 1262
QY 381 ProValLeuThrAlaLeuLeuTyArgLysIleLeuAspGluIleGluAlaAsnAsp 400
Db 1263 CCTGTATGGGCATCTTTGCTCTTGACCGCAAAATACTAGATGAGATTGAAGCCAATGAC 1322
QY 401 TyrAsnAspPheThrArgArgAlaTyValSerLysProLys-LysLeuLeuThrLeuPr 420
Db 1323 TACAACACTTCACAAGAGACATATGTGACAAATCAAGCAAGTGAATGCAATACC 1382
QY 420 oilealatyAlaLysSerLeuValProProAsnArgThrSerProLeuAlaLys 439
Db 1383 TATTGCATATGCAAAATCTTTGCTCCTCT-----ACAAACTGCCCTCTCTTCAA 1434

```

```

RESULT 12
LERYGTOM5
LOCUS LERYGTOM5 1355 bp mRNA linear PLN 02-AUG-1993
DEFINITION L.esculentum (ry mutant) GTOM5 mRNA for mutant phytoene synthase.
ACCESSION X67144
VERSION X67144.1 GI:19346
KEYWORDS GTOM5 gene; mutant; phytoene synthase.
SOURCE Lycopersicon esculentum.

```


LOCUS AB037975 1689 bp mRNA linear PLN 25-NOV-2000
 DEFINITION Citrus unshiu mRNA for phytoene synthase, complete cds.

AB037975
 VERSION AB037975.1 GI:11344506
 KEYWORDS phytoene synthase.
 SOURCE Citrus unshiu (cultivar:Satsuma mandarin) mature stage pulp cDNA to mRNA.

ORGANISM

Citrus unshiu
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.
 1 (sites)
 Ikoma,Y., Komatsu,A., Kita,M., Ogawa,K., Omura,M., Yano,M. and
 Moriguchi,T.

TITLE

Expression of a phytoene synthase gene and characteristic carotenoid
 accumulation during citrus fruit development
 2 (bases 1 to 1689)
 Physiol. Plantarum 111, 232-238 (2001)

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (04-FEB-2000) Takaya Moriguchi, National Institute of
 Fruit Tree Science, Department of Research Planning and
 Coordination; 2-1 Fujimoto, Tsukuba, Ibaraki 305-8605, Japan
 (E-mail:takaya@fruit.affrc.go.jp, Tel:81-298-38-6416,
 Fax:81-298-38-6437)

FEATURES

Location/Qualifiers

1..1689
 /organism="Citrus unshiu"
 /cultivar="Satsuma mandarin"
 /db_xref="taxon:55188"
 /tissue_type="pulp"
 /dev_stage="mature stage"
 179..1489
 /codon_start=1
 /product="phytoene synthase"
 /protein_id="BAB18514.1"
 /db_xref="GI:11344507"

CDS

/translation="MSVTLVWVSPNSLSCGFFVDSVREENRFLYSRFLYQHOTR
 TAVENSRQFNNSKORNSYPLDLDLRHPCSSGDLPEISCHWASTAGEVAMSESE
 MYTVNLKQALVKNQPSGVTVDLNDPDLALPDLISLSEAYDRGGEYAKTFY
 LGTLLSERRALAIWYRCRTDELVDGPNASHITPTALDRWSEDLFRGRPF
 MDLALSDTVTFPPVDIQPFDMIEGRMDLRKSRKFNEDLYLYCYVAGVGLMSV
 PVMGIAPDSQATSVYNAALAIANOLNITLDVGDARRGRVYLPDLELAAGLS
 DDDIAGEVITIKRNFEMKIKRAMEFDMENGVTSEASRPVWASLLYRLQILD
 EIEANDYNEFTRAYVSKAKKIALPIAYAKSLRPSRIYTSKA"

BASE COUNT 512 a 302 c 388 g 487 t

ORIGIN

Alignment Scores:

Pred. No.: 2-71e-131 Length: 1689
 Score: 1626.00 Matches: 327
 Percent Similarity: 82.61% Conservative: 34
 Best Local Similarity: 74.83% Mismatches: 64
 Query Match: 71.63% Indels: 12
 DB: 8 Gaps: 4

US-09-847-081b-2 (1-440) x AB037975 (1-1689)

Qy 2 SerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThr 21

Db 176 AGCATGCTGTACATTCGCTGGTGTATCACCCTAACCTACCAATGTCTCAATGGCTC 235

Qy 22 GlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerArgPheLeu--- 40

Db 236 GGGTCTCGATCAGTTCGAGAGAAACAGGCTGTTTATTCATCAAGATTTCTTTAC 295

Qy 41 ---AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLys-----GlyGly 56

Db 296 CAACATCAACCCGAGCTGCTGTGTTAAATCTTAGACCTAAGCAGATTAATAATAGPAAT 355

Qy 57 ArgGlnArgTrpAsnGlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGly 76

Db 356 AAGCAGACGAGGATTTCTTATCTCTTTAGATACAGATTTGAGGCATCTCTGC----- 406

Qy 77 SerArgThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAla 96

Db 407 -----TCATCTGGAAATCGACTTGCCTGAAATATCATGTATGTTGCTAGCAGCT 457

Qy 97 GlyGluMetThrValSerSerGluLysLysValValValValValValValValValVal 116

Db 458 GGNAGAGTGGCCATGCTCTCAGAGAAATGTTTACAATGTTGCTCAAGACGAGGACCC 517

Qy 117 LeuValLysArgGlnLeuArgSer---ThrAspAspLeuGluValLysProAspIleVal 135

Db 518 TTGGTTAATAAGCAACCAAGTGGGTACTCGTGATCTTCATGTGAACCCAGATATTGCT 577

Qy 136 ValProGlyAsnLeuGlyLeuLeuSerGluAlaValAspArgCysGlyGluValCysAla 155

Db 578 TTACCCGGAACTTAAAGTCTGCTCAGTGAAGCTTATGATCGTTGTTGAGAGATTTGGCC 637

Qy 156 GluTrpAlaLysThrPheTrpLeuGlyThrLysLeuMetThrProGluAlaArgAla 175

Db 638 GAGTATGCTAAGACATTTTACTTTGGGACCTTTCGTGATGACCTCTGAAGGCGAGGCT 697

Qy 176 IleTrpAlaLysThrValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAla 195

Db 698 ATATGGCTATATATGCTGCTGAGGACAGATGAGCTCGTTGATGGCCCTAATGCT 757

Qy 196 SerHisIleThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSer 215

Db 758 TCACACATTAACCAACAGCTTTAGACAGGTGGAGTCCAGGTGGAGACCTTTTCCGG 817

Qy 216 GlyArgProPheAspMetLeuAspAlaLeuSerAspThrValSerArgPheProVal 235

Db 818 GGTGCTCCATTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 877

Qy 236 AspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArg 255

Db 878 GACATTCAGCCATTCAGAGATATGATAGAAGGAATGAGGATGGACCTTAGGAAGCTCA 937

Qy 256 TyrLysThrPheAspGluLeuLeuTrpCysTrpTyrValAlaGlyThrValGlyLeu 275

Db 938 TACAAAACCTTTGATGAATATATCTGTTATTTATTTATTTATTTATTTATTTATTT 997

Qy 276 MetSerValProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyr 295

Db 998 ATAGTGTTCACATTATGGCATAGCACCTGACTCAGGCAACACAGAGAGGCTCTAC 1057

Qy 296 AsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGly 315

Db 1058 AATGCAGCATTTGGCATTGGGATTTGCTAATCAGCTCACTAACATACATCAGAGAT 1117

Qy 316 GluAspAlaArgArgGlyArgValTrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeu 335

Db 1118 GAGGATGCCGGAAGAGAGAGGTTTATCTACCAAGATGAGTTGGCAGCAGGAGGCTT 1177

Qy 336 SerAspGluAspIlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLys 355

Db 1178 TCAGATGATGACATATTTGCTGGAGAGTGAACCATTAATGGAGAACTTCATGAAGAAC 1237

Qy 356 GlnIleGlnArgAlaArgLysPheAspGluSerGluLysGlyValThrGluLeuAsp 375

Db 1238 CAAATTAAGAGGCAAGGATGTTCTTTGATGCTCAGAACGCTGTGACCGAGCTGAGT 1297

Qy 376 SerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTrpArgLysIleLeuAspGlu 395

Db 1298 GAAGCTAGTCGATGGCGGTATGGCTTCATTTGTTTACCGGCAAACTACTGGATGAG 1357

Qy 396 IleGluAlaAsnAspTrpAsnAsnPheThrArgArgAlaTrpValSerLysLysLys 415

Db 1358 ATTGAGGCCAATGATTACAACACTTCACAAAGAGAGCTTATGTGAGTAAGCAAGAAC 1417

Qy 416 LeuLeuThrLeuProIleAlaTrpAlaLysSerLeuValProAsnArg 432

Db 1418 ATAGCTCATTACCAATTTGATATGCAAAATCCCTCTTACCCCGCTCAAGA 1468

RESULT 15

AF152892 1773 bp mRNA linear PLN 04-APR-2001
 LOCUS Citrus x paradisi phytoene synthase mRNA, complete cds.
 DEFINITION AF152892
 ACCESSION AF152892
 VERSION AF152892.2 GI:13542331
 KEYWORDS
 SOURCE Citrus x paradisi.
 ORGANISM Citrus x paradisi

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.

REFERENCE

AUTHORS Costa, M.C., Moreira, C.D., Melton, J.R., Otoni, W.C. and Moore, G.A.
 TITLE Developmental expression of carotenoid genes in Citrus
 JOURNAL Unpublished
 REFERENCE 1 (bases 1 to 1773)
 2 (bases 1 to 1773)
 AUTHORS Costa, M.C., Moreira, C.D., Melton, J.R., Otoni, W.C. and Moore, G.A.
 TITLE Direct Submission
 JOURNAL Submitted (18-MAY-1999) Horticultural Sciences, University of
 Florida, 1301 Field Hall, Gainesville, FL 32611, USA
 REFERENCE 3 (bases 1 to 1773)
 AUTHORS Costa, M.C., Moreira, C.D., Melton, J.R., Otoni, W.C. and Moore, G.A.
 TITLE Direct Submission
 JOURNAL Submitted (04-APR-2001) Horticultural Sciences, University of
 Florida, 1301 Field Hall, Gainesville, FL 32611, USA
 REMARK Sequence update by submitter
 COMMENT On Apr 4, 2001 this sequence version replaced gi:5020351.
 FEATURES Location/Qualifiers

source

1..1773
 /organism="Citrus x paradisi"
 /db_xref="taxon:37656"
 251..1561
 /codon_start=1
 /product="phytoene synthase"
 /protein_id="RAD38051.2"
 /db_xref="GI:13542332"

CDS

/translation="MSVTLVWVSPNSOLSNCFGVDSVREENLFYSSRFYQHQR
 TAVNSRPFQNNKQRRNSYPLDLDLHPCSSGIDLPESCVAAPGAVMSSEE
 MYNVLLQAAVLNKPQSGVTRDLDVNPDIAPGLTSLSEAYDRCGVCAEYAKTFY
 LGTLMTSERRAIWAIYWCRRDLDVDPNASHITPTALDRWESRLDLFRGQPED
 MLDALSDTVKFPVDIQPFDMLEGHMDLRKSNKFNEDLYLYCYVAGTVGLMSV
 PMGTAPQSQATTSYVNAALGIANQLTNILRDVGEDAPRGVYLPQDELARGLS
 DDIFAGETVILKRNFMNQLKRAFMFDMENGVTSEASRMPVWASLLLYRQILD
 EIEANDYNNFTKRCVSKAKKIAALPIAYAKSLRPSRIYTSKA"
 BASE COUNT 536 a 319 c 395 g 523 t
 ORIGIN

Alignment Scores:

Pred. No.: 1..55e-129 Length: 1773
 Score: 1606.00 Matches: 324
 Percent Similarity: 82.1% Conservative: 35
 Best Local Similarity: 74.14% Mismatches: 66
 Query Match: 70.75% Indels: 12
 DB: 8 Gaps: 4

US-09-847-081B-2 (1-440) x AF152892 (1-1773)

QY 2 SerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThr 21
 DB 248 AGCATGCTGTACATGCTGTGGGTGTATCACCTAACTCAATGTTCCAATGCTTC 307
 QY 22 GlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu--- 40
 DB 308 GGGTTCGTGATCAGTTCGAGAGGAAACAGCGCTTTTATTCATCAGAGATTCCTTAC 367
 QY 41 ---AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLys-----GlyGly 56
 DB 368 CAACATCAACACCGCGACTGCTGTGTGTTAATCTACACCTTAAGCAGTTAATAATACTAAT 427
 QY 57 ArgGlnArgTrpAsnPhcGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGly 76
 DB 428 AAGCAGAGAGGAGNATCTTATCTCTTAGATACAGATTTCAGGCATCCTTCG----- 478

QY 77 SerArgThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAla 96
 DB 479 -----TCATCTGGAAATCAGCTTGCCTGAAATATCATGTATGTTGCTACACTGCT 529
 QY 97 GlyGluMetThrValSerSerGluLysLysValTyrAspValValLeuLysGlnAlaAla 116
 DB 530 GGAGAAGTGGCCATGCTTTCAGAGAATAAGTGTTCACATGTTGCTCAAGCAGGAGCC 589
 QY 117 LeuValLysArgGlnLeuArgSer---ThrAspAspLeuGluValLysProAspIleVal 135
 DB 590 TTGGTTAATAACCAACCAAGTGGGTACTCGTATGCTTTCATGATGAACCCAGATATGCT 649
 QY 136 ValProGlyAsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAla 155
 DB 650 TTACCCGGGAACCTTAAAGTCGCTCAGTGAAGCTTATGATCGTTGTCGAGAGTTTCGCC 709
 QY 156 GluTyrAlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAla 175
 DB 710 GAGTATGCTAAGACATTTTACTTGGAACTTTGCTGATGACCTCTGAAAGGCCAAGGCT 769
 QY 176 IleTyrAlaLysThrValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAla 195
 DB 770 ATATGGGCTATATATGTGTGTAGGAGGACAGATGAGTCTGATGGGCTAATGCT 829
 QY 196 SerHisIleThrProGlnAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSer 215
 DB 830 TCACACATAACTCCACACACTTTAGACAGTGGGAGTCCAGGTTGGAAGACCTTTCCCG 889
 QY 216 GlyArgProPheAspMetLeuAspAlaLeuSerAspThrValSerArgPheProVal 235
 DB 890 GGTACGCCATTTGATATGCTTGCATTCATATCAGATACAGTAACCAAAATTCCTGTC 949
 QY 236 AspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArg 255
 DB 950 GACATTCAGCCATTCAGAGATATGATAGAAGGAATGAGGATGACCTTAGGAAGTCAAGA 1009
 QY 256 TyrLysThrPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeu 275
 DB 1010 TACAAAACCTTTGATGAATATATCTGTTATGTTATGTTGCTGGACCGTAGGCTA 1069
 QY 276 MetSerValProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyr 295
 DB 1070 ATGAGTGTTCAGTATGTCGATAGCACCTGACTCAGCGCAACACAGAGAGCGCTAC 1129
 QY 296 AsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGly 315
 DB 1130 AATGAGCATTTGGCACTAGGATGCTAATCAGCTCCTAATCAGTCTCAGAGATGTTGA 1189
 QY 316 GluAspAlaArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeu 335
 DB 1190 GAGGATGCCCAAGAGGAGGTTTATCTACACAGATGAGTTCGCACAGGCGAGGCTT 1249
 QY 336 SerAspGluAspIlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLys 355
 DB 1250 TCAGATGATGACATATTTGCTGGAGAGGTGACCATTAATGGAGAACTTCATGAAGAAC 1309
 QY 356 GlnIleGlnArgAlaArgLysPheAspGluSerGluLysGlyValThrGluLeuAsp 375
 DB 1310 CAATTAAGAGGCGAAGGATGTTCTTTGATATGGCTGAGAACCGGTGACCGAGCTAGT 1369
 QY 376 SerAlaSerArgTrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGlu 395
 DB 1370 GAAGTAGTCGATGGCGGCTATGGGCTTCATTGCTTGTACCGGCAATATCTGGATGAG 1429
 QY 396 IleGluAlaAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLys 415
 DB 1430 ATTGAGGCCAATGATTACAACTTTCACAAAGAGAGCTGTGTGAGTAAGCAAGAAG 1489
 QY 416 LeuLeuThrLeuProIleAlaTyrAlaLysSerLeuValProProAsnArg 432
 DB 1490 ATAGCTGCACCTACCAATTCATATGCAAAATCCCTCTTACGCCCGTCAAGA 1540

Search completed: April 5, 2003, 04:01:14
Job time : 3111 secs
